Three-state protein folding: Experimental determination of free-energy profile

EKATERINA N. BARYSHNIKOVA, BOGDAN S. MELNIK, ALEXEI V. FINKELSTEIN, GENNADY V. SEMISOTNOV, AND VALENTINA E. BYCHKOVA

Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

(RECEIVED February 9, 2005; FINAL REVISION June 18, 2005; ACCEPTED July 6, 2005)

Abstract

When considering protein folding with a transient intermediate, a difficulty arises as to determination of the rates of separate transitions. Here we overcome this problem, using the kinetic studies of the unfolding/refolding reactions of the three-state protein apomyoglobin as a model. Amplitudes of the protein refolding kinetic burst phase corresponding to the transition from the unfolded (U) to intermediate (I) state, that occurs prior to the native state (N) formation, allow us to estimate relative populations of the rapidly converting states at various final urea concentrations. On the basis of these proportions, a complicated experimental chevron plot has been deconvolved into the urea-dependent rates of the I↔N and U↔N transitions to give the dependence of free energies of the main transition state and of all three (N, I, and U) stable states on urea concentration.

Keywords: protein folding; folding intermediates; tryptophan fluorescence; chevron plot; stopped-flow; apomyoglobin

Apomyoglobin is a good object for protein folding studies because its thermodynamic (Griko et al. 1988; Hughson et al. 1990; Jennings and Wright 1993; Jamin et al. 2000) and structural (Barrick and Baldwin 1993a,b; Eliezer and Wright 1996; Eliezer et al. 1998; Jamin and Baldwin 1998; Jamin et al. 1999; Lecomte et al. 1999; Tcherkasskaya and Ptitsyn 1999; Tsui et al. 1999; Tcherkasskaya et al. 2000) properties in both the native and intermediate conformational states are well elucidated. At neutral pH, apomyoglobin structure is "native" globular, with seven of eight helices of holomyoglobin tightly packed (A–E and G, H; while F, involved in the heme binding, is disordered in apoprotein) (Eliezer and Wright 1996). At pH 4.2, the native structure undergoes the transition into a thermodynamically stable "molten globule" (Dolgikh et al. 1981; Ptitsyn 1995) intermediate state (Griko et al. 1988) that contains three helices (Hughson et al. 1990), A, G, and H. This intermediate has two sub-states (stable at pH 4.2 and pH 3.9), which convert one into the other within a millisecond time range (Jamin and Baldwin 1998; Jamin et al. 1999). It was shown using the stopped-flow and quench-flow techniques that urea-induced apomyoglobin refolding goes via a kinetic intermediate that forms within 6 msec and is structurally similar to the equilibrium molten globule intermediate observed at pH 4.2 (Jennings and Wright 1993). Subsequent kinetic studies suggested that this intermediate is on-pathway (Jamin and Baldwin 1998). Quench-flow amide proton exchange combined with mass-spectrometry confirmed that apomyoglobin folds by a single pathway and that the intermediate is obligatory (Tsui et al. 1999). NMR analysis of mutant apomyoglobins also showed that some point mutation may change the folding pathway of apomyoglobin (Garcia et al. 2000).

The presence of a kinetic intermediate on the apomyoglobin folding pathway makes this protein attractive for studies of three-state folding/unfolding reactions. Protein folding involves a transition state and, for the majority of proteins, kinetic intermediates whose structural and

Reprint requests to: Valentina E. Bychkova, Institute of Protein Research (Moscow office), Room 104, Vavilova Street 34, Moscow, GSP 1, 117334, Russia; e-mail: bychkova@vega.protres.ru; fax: +7-095-135-9984.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.051402705.

thermodynamic properties are essential for the understanding of the protein folding mechanism. However, the study of these states, i.e., kinetic intermediates and transition states, is hampered by two factors. First, early folding intermediates are often formed within 10^{-3} sec (Schreiber and Fersht 1993; Munoz et al. 1994; Kay and Baldwin 1996; Parker et al. 1997; Burns et al. 1998; Cavagnero et al. 1999; Parker and Marqusee 1999; Tang et al. 1999; Laurents et al. 2000); these are relatively compact and have a secondary structure but no tight packing of side chains (Kim and Baldwin 1990; Clarke and Waltho 1997; Roder and Colon 1997), and, therefore, the rate of their formation cannot be measured in stopped-flow experiments. Second, the rates obtained from kinetic unfolding/refolding experiments often depend on a combination of fast and slow processes. Hence, there arises a problem of distinguishing between rate constants of separate events. Besides, structural studies of transition states require knowledge of the rate of crossing each separate energy barrier over the entire range of denaturant concentrations; this only allows applying the Φ-analysis that is currently the main experimental approach used to outline the structure of transition states (Matouschek et al. 1990, 1998; Itzhaki et al. 1995).

In this paper, we use the amplitude of the burst (U↔I) phase of apomyoglobin refolding to estimate the population of the kinetic intermediate (I) in various conditions. This allows us to estimate the rates of individual I→N and N→I transitions (which is the rate-limiting step of apomyoglobin folding or unfolding, respectively). As a result, the relative positions of free energy levels for the native, intermediate, unfolded, and the rate-limiting transition state (TS) are evaluated for apomyoglobin at various urea concentrations. This approach can be useful in studying factors affecting stability of the N, I, U, and TS states, as well as in structural characterization of TS and I states by protein engineering using directed mutations (Fersht et al. 1992).

Results

Fluorescence-detected intermediate state is accumulated upon urea-induced equilibrium unfolding of apomyoglobin

We studied equilibrium unfolding of apomyoglobin at 11°C by far UV CD and found that it is well described by a two-state transition with the mid-point at 2.9 M urea (Fig. 1). Previous far UV CD studies of urea-induced equilibrium denaturation of apomyoglobin at other temperatures also revealed no noticeable accumulation of thermodynamically stable intermediates (Cavagnero et al. 1999). However, changes in the protein tryptophan (Trp) fluorescence spectrum that accompany increasing urea concentration unequivocally reveal accumulation of some stable or metastable equilibrium intermediate states, whose Trp fluorescence is noticeably higher than that of the native (N) and unfolded (U) conformations (Fig. 2). Besides, this alteration of the Trp fluorescence spectrum is characterized by two isosbestic points at 315 and 370 nm that also confirm the presence of at least two transitions detectable by Trp fluorescence (Fig. 2).

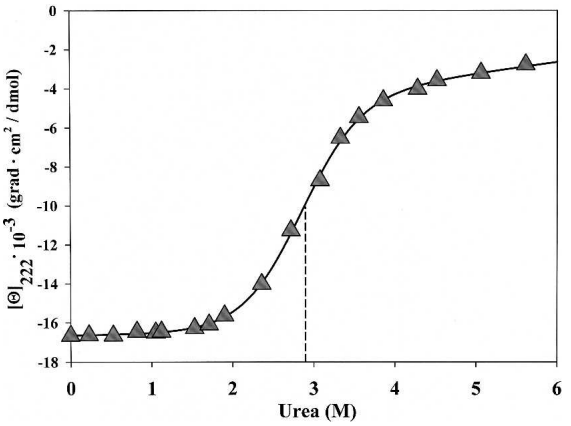

Figure 1. Urea-induced equilibrium unfolding of sperm whale apomyoglobin monitored by intensity of the far UV CD spectra at 222 nm. The dashed line indicates the mid-transition point. The solid line represents an approximation of the transition by the two-state model (Santoro and Bolen 1988).

For apomyoglobin refolding/unfolding kinetics experiments a wavelength of 335 nm has been chosen. As shown in Figure 2, the Trp fluorescence intensity change at this wavelength reflects accumulation of both an intermediate and the unfolded state. Interestingly, the I→U transition registered by fluorescence at 335 nm is more pronounced than that registered by the maximum intensity (Fig. 2, inset). The unfolded state displays a much lower intensity at 335 nm than the intermediate and native states, while the maximum intensity of the unfolded state is less than that of the intermediate state but exceeds the intensity of the native state. This is underlain by an additional decrease in 335 nm intensity caused by a shift of the fluorescence spectrum towards longer wavelengths in the unfolded state (Fig. 2). Thus, the fluorescence intensity at 335 nm proves to be more informative as to changes in populations of the native, intermediate, and unfolded states during kinetic unfolding/refolding reactions of apomyoglobin, and it was used in kinetic measurements.

Apomyoglobin refolding from the urea-unfolded state goes via a kinetic intermediate with high fluorescence intensity

We performed kinetic experiments on apomyoglobin unfolding and refolding monitored by Trp fluorescence

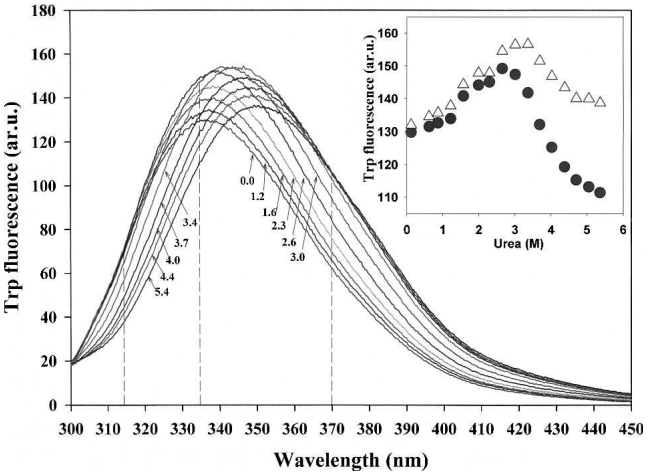

Figure 2. Trp fluorescence spectra of apomyoglobin at various urea concentrations (shown by arrows and numbers near curves). Dashed lines indicate 335 nm wavelength and the pseudoisosbestic points at 315 and 370 nm. (*Inset*) Dependence of protein Trp fluorescence intensities at 335 nm (filled circles) and the spectrum maximum intensity (open triangles) on urea concentration.

at 335 nm. Figure 3 presents time-resolved courses of the Trp fluorescence changing during apomyoglobin refolding (from 5 M urea to various final urea concentrations). At final urea concentrations below 3 M, there are two consecutive refolding phases: The first phase occurs within the dead time of a stopped-flow instrument and is revealed by a jump-wise increase of fluorescence intensity, and the second phase is observed as a slow decrease of fluorescence intensity. At final urea concentrations above 3 M, there is only one fast phase, which manifests itself as a burst-like insignificant increase of fluorescence intensity. So, owing to the instrument dead time, it is only the result of the fast phase (i.e., the transition from the unfolded to kinetic intermediate state of apomyoglobin) that can be observed. After protein refolding is completed (i.e., time$\rightarrow\infty$) the fluorescence intensity values correspond to equilibrium values. Figure 3 (inset) shows dependence of these values on urea concentration. One can see that the curve for equilibrium urea-induced transition coincides with that for $I_{335(\text{time}\rightarrow\infty)}$ obtained from protein refolding kinetics at a final urea concentration. Thus, we can assume that accumulation of the equilibrium intermediate state during urea-induced unfolding of apomyoglobin (Fig. 2) is a result of interconversion between the native, intermediate, and unfolded state populations at increasing urea concentration. It should be noted that the kinetic intermediate (like the equilibrium intermediate shown in Fig. 2) has a higher fluorescence intensity (at 335 nm) than that of the native or unfolded state. This property of the intermediate state is used to separate the kinetic transition U$\leftrightarrow$I from the transition I$\leftrightarrow$N. Since the slow phase of apomyoglobin refolding always leads to a decrease of fluorescence intensity, folding into the native state is believed to start from the intermediate state. At a given urea concentration M, the transient intermediate state population $f_I(M)$ can be calculated from the burst phase amplitude A(M) (see Fig. 4) according to the equation:

$$f_I(M)=\frac{\left(A(M)-A_U(M)\right)}{\left(A_I(M)-A_U(M)\right)}. \quad (1)$$

Here $A_I(M)$ and $A_U(M)$ denote the intermediate and unfolded state amplitudes. Their dependence on urea concentration M is derived from extrapolation of the pre- and post-transition baselines into the transition region (Fig. 4, inset).

This population $f_I(M)$ is large at small final urea concentrations (Fig. 4) and reaches its maximal level at 1.2 M urea, which means that at this (and lower) urea

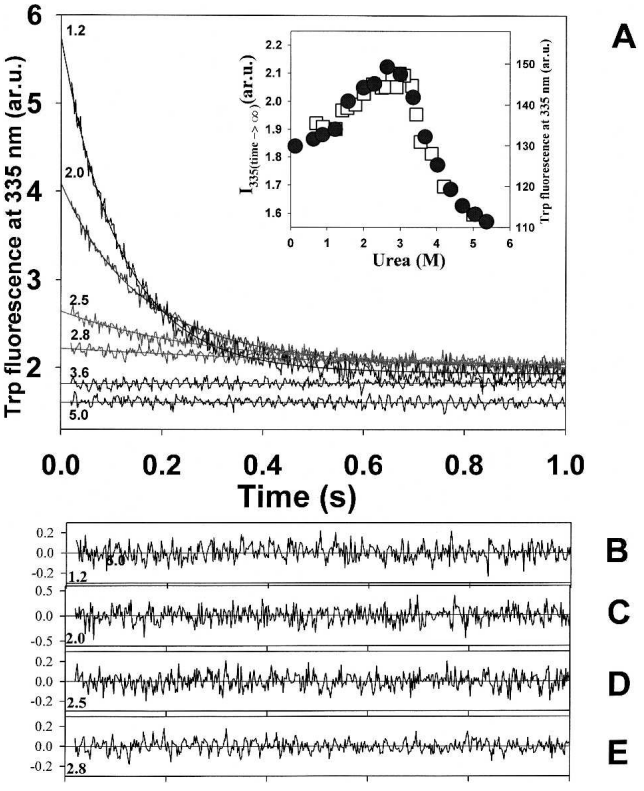

Figure 3. Apomyoglobin refolding experiments. (*A*) Refolding kinetics monitored by Trp fluorescence at 335 nm. The initial urea concentration was 5.0 M; numbers near the curves indicate final urea concentrations. Solid lines represent single-exponential approximations of the kinetics to zero time. (*B–E*) Residuals between experimental curves and those calculated from a single-exponential fit; this fit gives the observed rate (k_{obs}) of coming to equilibrium. (*Inset*) Urea-induced dependence of protein fluorescence intensities at 335 nm (filled circles) at equilibrium unfolding and of $I_{335(\text{time}\rightarrow\infty)}$ (open squares) obtained from protein refolding kinetics.

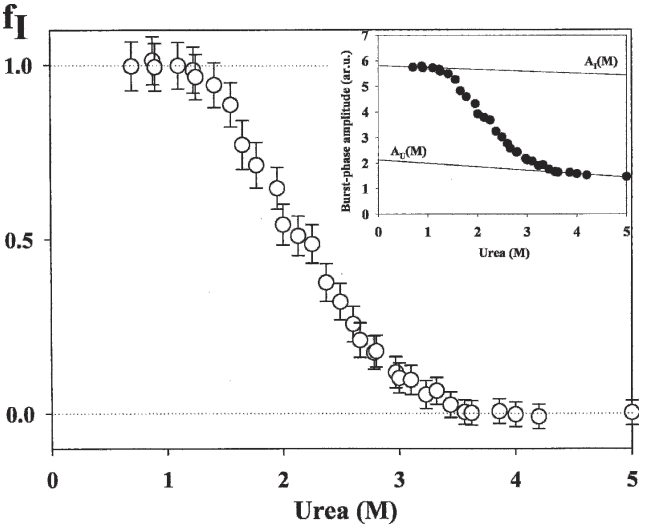

Figure 4. Population (f_I) of the kinetic intermediate calculated from the burst phase amplitude according to Equation 1. (*Inset*) Dependence of the burst phase refolding amplitude A(M) on final urea concentration. A_U(M) and A_I(M) are fluorescence intensities for the unfolded and intermediate state, respectively.

concentration, all protein molecules start their transition into the N state after they have acquired the intermediate state structure. The 2.1-M urea concentration provides a 50% intermediate state population, which means that stability of the intermediate state is equal to that of the unfolded state.

Unlike population of the *kinetic* I state at folding (which achieves its maximum at a urea concentration below 1.2 M, see Fig. 4), the equilibrium population of the I state is maximal at 2.8 M urea (Fig. 2). Therefore, at 2.8 M urea, the I state at equilibrium is 5–10 times less populated than the U state (Fig. 4); that, in turn, is a little less populated (at equilibrium) than the N state. A decrease of urea concentration below 2.8 M makes the I state even more unstable (and, therefore, less populated at equilibrium) than the N state. An increase of urea concentration above 2.8 M makes the I state very unstable relative to the U state (and, therefore, progressively less populated than the U state both in kinetics and at equilibrium).

Urea-induced apomyoglobin unfolding goes via a kinetic intermediate

Figure 5 shows kinetic curves for apomyoglobin unfolding measured by fluorescence intensity at 335 nm for various final urea concentrations. Unlike folding, the unfolding has no burst phase and shows up as a comparatively slow change of fluorescence intensity. Up to the 3.5 M final urea concentration, the intensity increases, but higher values of urea concentration are accompanied by its decrease. This interesting phenomenon can be interpreted as follows. The native state unfolds to a mixture of the I and U states, whose equilibrium is achieved almost instantly. The I state increases the level of Trp fluorescence intensity, while the U state decreases this level, as compared to that of the N state (Fig. 3). At a high I/U ratio, the unfolding demonstrates an increasing total fluorescence, but with a minor contribution of the I state, i.e., at a final urea concentration above 3.5 M, the total fluorescence decreases.

The unfolding kinetics rate is determined predominantly by overcoming a large free energy barrier between the N and I (and hence U) states. The subsequent fast I↔U conversion gives virtually no contribution to the rate of either N→I (k_{NI}) or N→U (k_{NU}) processes, and, hence, $k_{NI} = k_{NU}$. This allows us to make a linear extrapolation of ln ($k_{NI} = k_{NU}$) to the region of mid-transition and further to the region of refolding urea concentrations, which is to be used in estimating folding rates.

Discussion

Analysis of the chevron plot

Both folding and unfolding bring populations of different protein states to equilibrium. The resulting populations

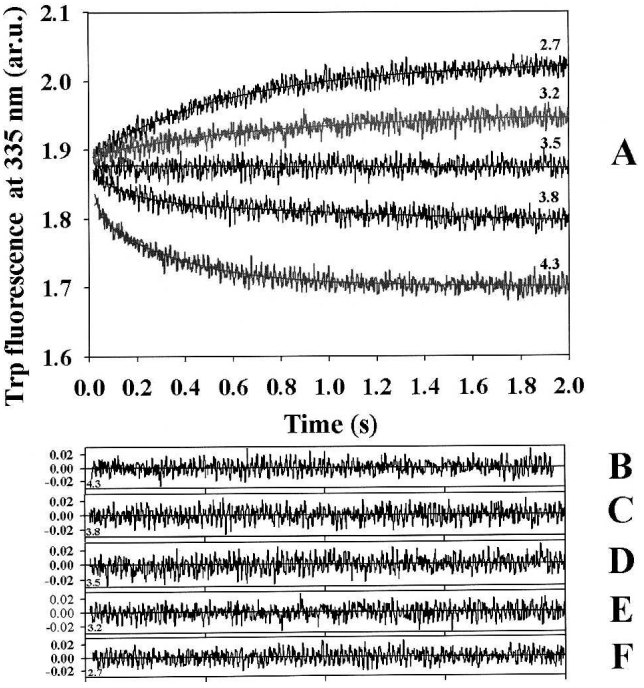

Figure 5. Apomyoglobin unfolding experiments. (*A*) Unfolding kinetics monitored by Trp fluorescence at 335 nm. The initial urea concentration was 0.0 M; numbers near the curves indicate final urea concentrations. (*B–F*) Residuals between experimental curves and those calculated from a single-exponential fit.

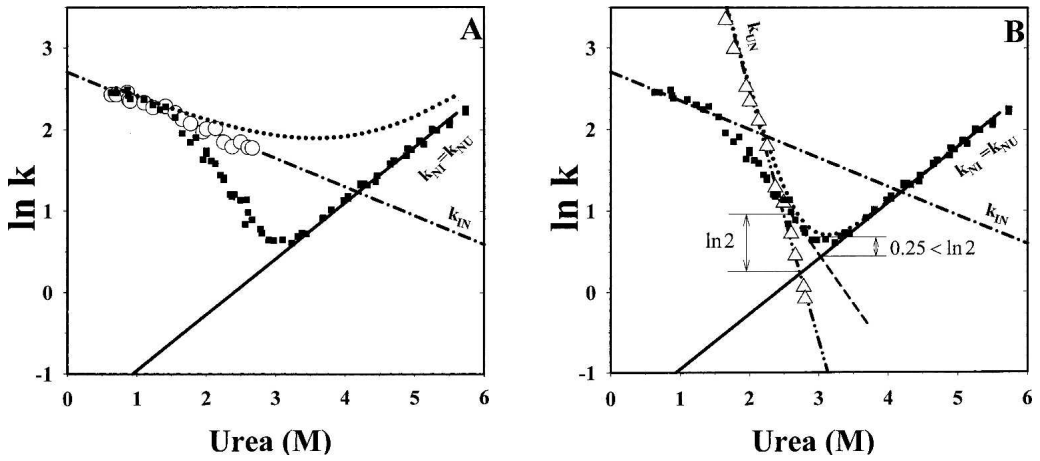

Figure 6. An experimental chevron plot for the observed apomyoglobin folding/unfolding rate constants (k_{obs}) and rate constant estimates for elementary transitions (k_{IN}, k_{NI}, k_{UN}, k_{NU}). Filled squares show the observed rates (k_{obs}) of coming to equilibrium at indicated urea concentrations. The left branch of the chevron plot (<3 M urea) corresponds to folding, the right branch (>3 M urea) to unfolding. (*A*) Estimation of the I→N transition rate (k_{IN}). Open circles indicate ln(k_{IN}) values calculated from Equation 3 using k_{obs} of the chevron folding branch, extrapolated k_{obs} of the unfolding branch (solid line $k_{NI} = k_{NU}$), and the kinetic intermediate population $f_I(M)$ calculated from the fast refolding phase amplitude (see Fig. 4). The dash-dotted line represents extrapolation of k_{IN} to the region of higher urea concentrations. The dotted line shows the estimated rate of coming to equilibrium for a separate I↔N transition. (*B*) Estimation of the U→N transition rate, k_{UN} (the dash-dot-dotted line). The estimates (open triangles) are calculated according to Equation 6 using k_{obs} values, the $k_{NI} = k_{NU}$ line, and the $f_I(M)$ curve (see Fig. 4). The dash-dotted and solid lines denote the same as in Figure 6A. The dashed line represents a simple linear extrapolation of experimental points (k_{obs}) of the folding branch to the mid-transition region. The right arrow shows a difference between the point of intersection of the extrapolated folding and unfolding branches and the experimental points at the same urea concentration; this difference (≈0.25) is too small to be compatible with that expected for a two-state transition (ln 2). However, the difference between the point of intersection of ln k_{UN} and ln k_{NU} and the chevron plot for a separate U↔N transition (dotted line) is close to ln 2 (left arrow).

depend on denaturant concentration, i.e., the native state dominates at low final concentrations of urea, while the unfolded state dominates at its high final concentrations. By plotting the logarithm of the observed rate of coming to equilibrium (ln k_{obs}) as a function of final urea concentration, we obtained the so-called chevron plot (Fig. 6).

Since apomyoglobin unfolding, like the rate-limiting phase of its folding, is well described by single-exponential kinetics (see Figs. 3, 5), one might think that the obtained chevron plot describes a two-state folding. However, there is a pronounced difference between our plot (Fig. 6B) and chevrons typical of two-state folding proteins (e.g., CI2, Itzhaki et al. 1995). First, in the region of low urea concentration (<1.5 M), there is a rollover of the folding limb, which is usually believed to originate either from admixture of an intermediate state (Matouschek et al. 1990) or from a change within a transition state (Ternstorm et al. 1999). Second, the rate of transition in the region of the chevron plot's minimum seems to be too low, i.e., too close to the point of intersection of extrapolated folding and unfolding limbs of the chevron. Indeed, the observed rate of coming to equilibrium in the case of a two-state reaction is the sum of rate constants for protein folding and unfolding ($k_{obs} = k_f + k_u$) (Fersht 2000). At the mid-transition point $k_f = k_u$, i.e., $k_{obs} = 2 \cdot k_u$, and ln (k_{obs}) = ln (2) + ln (k_u). Hence, the ln (k_{obs}) minimum should be by ln (2) ≈ 0.7 higher than the intersection of the folding and unfolding limbs. In our case, however, this value is about 0.25 (Fig. 6B). Taking into account that rate constants of transition over the rate-limiting free energy barrier between I and N (k_{NI} for the N→I transition and k_{IN} for I→N) are considerably less than the I↔U rates k_{UI} and k_{IU} (here, about 1 sec^{-1} and 10^{-3} sec^{-1}, respectively), the experimentally observed rate constants can be described by the equation (Parker et al. 1995; Baldwin 1996)

$$k_{obs} = k_{NI} + f_I \cdot k_{IN}, \quad (2)$$

where f_I is an intermediate state population proportional to the amplitude of the burst phase of refolding kinetics (Fig. 3). Thus, the k_{IN} value, at given urea concentration (M), can be obtained as

$$k_{IN}(M) = \frac{k_{obs}(M) - k_{NI}(M)}{f_I(M)}, \quad (3)$$

where f_I, k_{obs}, and $k_{NI} = k_{NU}$ (or its extrapolated value) are taken from the experimental data (Figs. 4, 6A). This estimate is reliable in folding conditions [when $k_{obs}(M)$ is sufficiently large as compared to $k_{NI}(M)$], but only if $f_I(M)$ is not too small to be reliably estimated (according to Fig. 4, this requirement is satisfied when the final urea concentration is below 2.6–2.7 M). The most reliable estimate of k_{IN} can be obtained when $f_I(M) \approx 1$. As seen from Figure 4, $f_I(M) \approx 1$ occurs at final urea concentrations below 1.2 M, when all protein molecules are in the intermediate state just after the burst phase. Hence, within this range of urea concentrations, the unfolded state is much less stable than the I state and does not affect the rate of I→N transition at all. However, this linear part of the chevron plot (corresponding to $f_I = 1$) is too small for apomyoglobin (Fig. 6A), and it is virtually absent for many of its mutants (E.N. Baryshnikova, B.S. Melnik, A.V. Finkelstein, G.V. Semisotnov, and V.E. Bychkova, unpubl.). To estimate k_{IN} and its dependence on urea concentration more accurately, we used the known values of $f_I(M)$, $k_{NI}(M)$, $k_{obs}(M)$, and Equation 3. The result is shown in Figure 6A (open circles). Then we approximated these points by a straight line and extrapolated it to the region of higher urea concentrations (Fig. 6A). Its intersection with the unfolding limb of the chevron plot corresponds to a 4.2 M urea concentration, where free energies of the native and intermediate states are equal. Using the rate constants k_{IN} calculated for the entire urea concentration range, we can construct a chevron plot for the rate-limiting free energy barrier between the I and N state (Fig. 6A).

The observed rate constant of protein folding is close to k_{IN} when f_I is close to 1, and the rate-limiting step of folding is determined only by the I→N transition, i.e., by the jump over the main free energy barrier starting from the I state. The height of the barrier is, in this case, $G_{TS} - G_I$, where G_{TS} and G_I are free energies of the transition (TS) and intermediate (I) state, respectively. However, at a higher (above 1.2 M) final urea concentration, the I state is less populated, and therefore, k_{obs} deviates from k_{IN} significantly, although in all cases the rate-limiting step of folding is determined by overcoming the same TS. The height of this barrier is $G_{TS} - G_U$ (where G_U is the U state free energy), and the U→N rate constant can be represented as $k_{UN} \sim \exp[-(G_{TS} - G_U)/RT]$ (where R is the gas constant and T is the temperature), while the I→N rate constant is $k_{IN} \sim \exp[-(G_{TS} - G_I)/RT]$. Hence, k_{UN} is equal to $k_{IN} \cdot \exp[-(G_I - G_U)/RT]$. Since $\exp[-G_I/RT]$ and $\exp[-G_U/RT]$ are proportional to populations of the I and U state (recall that the equilibrium between them is achieved instantly),

$$k_{UN} = k_{IN} \cdot \frac{f_I}{1 - f_I}. \quad (4)$$

Thus, Equation 2 can be written as

$$k_{obs} = k_{NI} + (1 - f_I) \cdot k_{UN} \quad (5)$$

[or, taking into account that $k_{NI} = k_{NU}$, as $k_{obs} = k_{NU} + (1 - f_I)\, k_{UN}$]. Hence, the k_{UN} value at a given urea concentration (M) can be calculated from

$$k_{UN}(M) = \frac{k_{obs}(M) - k_{NU}(M)}{1 - f_I(M)}. \quad (6)$$

This estimate can be reliable in folding conditions, but only if $1 - f_I(M)$ is not too small to be reliably determined (Figs. 4, 6A), i.e., when the final urea concentration is between 3.0 M and 1.5 M.

The result of calculation of the k_{UN} value is shown in Figure 6B (open triangles) together with extrapolation of $k_{UN}(M)$ to higher and lower urea concentrations. The rates of N→U and U→N transitions (lines k_{NU} and k_{UN} in Fig. 6B) are equal at 2.7 M urea, while the intersection of $k_{IN}(M)$ and $k_{UN}(M)$ lines occurs at 2.2 M urea, with $f_I = 0.5$ according to Figure 4.

Analysis of the three-state chevron plot for the case of unachieved rollover

The above described approaches allow one to estimate the rates of individual phases for a three-state kinetics folding. The apomyoglobin case is relatively simple in this respect because the chevron's rollover is evident. Can we perform a similar deconvolution of the rates of individual cases when the rollover is not observed? We considered this for an artificially created case when the intermediate state is poorly populated during the observed folding. This situation was modeled by cutting off A(M) and the chevron plot at 2.1 M urea (at this urea concentration, the apomyoglobin intermediate state is about 50% populated) (Fig. 7). Because after the burst phase there are only the unfolded and intermediate states, we approximated A(M) by a two-state model (Santoro and Bolen 1988). Then all the above mentioned estimates were made again. As seen from Figure 7, even if the observed population of the intermediate does not exceed 50%, we can estimate k_{IN} and k_{UN} with a reasonable accuracy. To do so, we used the SIGMA PLOT program to build up the observed 50% left limb to its 100%.

Plotting of the free-energy profile

The above results allow us to conclude that apomyoglobin refolding kinetics looks like a transition from the unfolded to an intermediate and further to the native state. An attempt to estimate the rate constants of

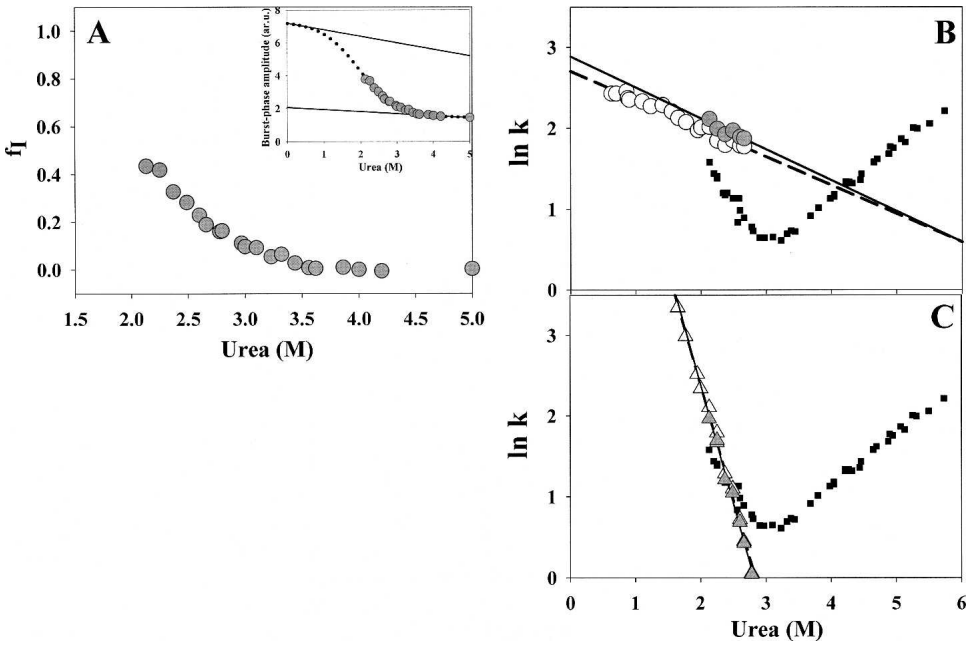

Figure 7. Estimation of the rate constants for the case in which the intermediate state does not reach 100% population during the observed folding reaction and when $f_I(M)$ cannot be obtained as described above. It was modeled by cutting off A(M) and the chevron plot at 2.1 M urea (shown in gray). Filled squares, open circles, and open triangles denote the same as in Figure 6. (*A*) Population (f_I) of the kinetic intermediate calculated from the burst phase amplitude according to Equation 1. (*Inset*) Dependence of the burst phase refolding amplitude A(M) on final urea concentration. The dotted line represents an approximation of A(M) by the two-state model. (*B*) Estimate for k_{IN}. The solid line represents extrapolation of k_{IN} to high urea concentrations. The dashed line denotes the same as the dash-dotted line in Figure 6. (*C*) Estimate for k_{UN}.

apomyoglobin transition between the native and intermediate states (k_{IN} and k_{NI}) was previously reported (Cavagnero et al. 1999), but it was restricted to extrapolation to zero denaturant concentration and did not allow constructing the chevron plot for the I↔N transition. The rate constants of the fast I↔U transition (k_{UI} and k_{IU}) cannot be measured experimentally because this event occurs within the stopped-flow dead time. Nevertheless, relative positions of free energies for the U, I, N, and TS states (G_U, G_I, G_N, and G_{TS}, respectively) can be estimated over the entire range of urea concentrations using the experimentally measured rate constants for protein folding/unfolding (k_{obs}) and percentage of population of the I state (f_I).

First, $G_I - G_U$ can be obtained, at various urea concentrations M, as

$$G_I(M) - G_U(M) = -RT \cdot \ln \frac{f_I(M)}{1 - f_I(M)}. \qquad (7)$$

Second, $G_N - G_I$ can be obtained from the N↔I two-state transition (Fersht 2000) as

$$G_N(M) - G_I(M) = -RT \cdot \ln \frac{k_{IN}(M)}{k_{NI}(M)}, \qquad (8)$$

where $k_{NI}(M)$ is the unfolding rate extrapolated to urea concentration M, and $k_{IN}(M)$ is the refolding rate determined from Equation 3.

Since it is convenient to count off all free energies from that of the unfolded state, Equations 7 and 8 can be combined to give:

$$\begin{aligned} G_N(M) - G_U(M) &= -RT \cdot \left[\ln \frac{k_{IN}(M)}{k_{NI}(M)} + \ln \frac{f_I(M)}{1 - f_I(M)} \right] \\ &= -RT \cdot \ln \frac{k_{UN}(M)}{k_{NU}(M)}. \end{aligned} \qquad (9)$$

In this transformation we used Equation 4 and equalities of k_{NI} and k_{NU}.

In its turn, the rate constant k_{NI} is determined by the $G_{TS} - G_N$ difference (see the transition state theory, Fersht 2000)

$$k_{NI} = \kappa \frac{RT}{h} \cdot \exp\left(-(G_{TS} - G_N)/RT\right), \qquad (10)$$

where h is the Planck constant and κ is the transmission coefficient that can be assumed to be equal to 1.0 (Fersht 2000, Chapter 18). Thus,

$$G_{TS}(M) - G_N(M) = -RT \cdot \left[\ln[k_{NI}(M)] - \ln\left(\kappa \frac{RT}{h}\right)\right] \quad (11)$$

and using Equation 9 we have

$$G_{TS}(M) - G_U(M) = -RT \cdot \left[\ln[k_{UN}(M)] - \ln\left(\kappa \frac{RT}{h}\right)\right] \quad (12)$$

where $RT \cdot \ln \frac{RT}{h} = 69.4 kJ/mol$ is the constant at 11°C (284K).

The free energy levels of all states for apomyoglobin folding/unfolding reactions are presented graphically in Figure 8. At low urea concentrations (below 1.2 M), both I and N states are more stable than the unfolded state, and the observed protein refolding kinetics corresponds to the I→N transition (Fig. 8). At higher urea concentrations (above 2.1 M), the intermediate state becomes less stable than the unfolded state. Fast redistribution of protein molecules between U and I states results in a relatively low population of the I state, and, hence, in a decreased probability of the I→N transition. This leads to a decrease of the observed refolding rate that becomes dependent on the direct U→N transition. Up to 2.7 M urea, the native state is still more stable then U (and I) and, therefore, the protein folds. At urea concentration above 2.7 M, unfolding of the protein becomes more favorable. Up to 4.2 M urea, the native state is more stable than the intermediate state (Figs. 6, 8); above this urea level, the intermediate is more stable than the native state, but both these states are strongly destabilized in comparison with the unfolded one.

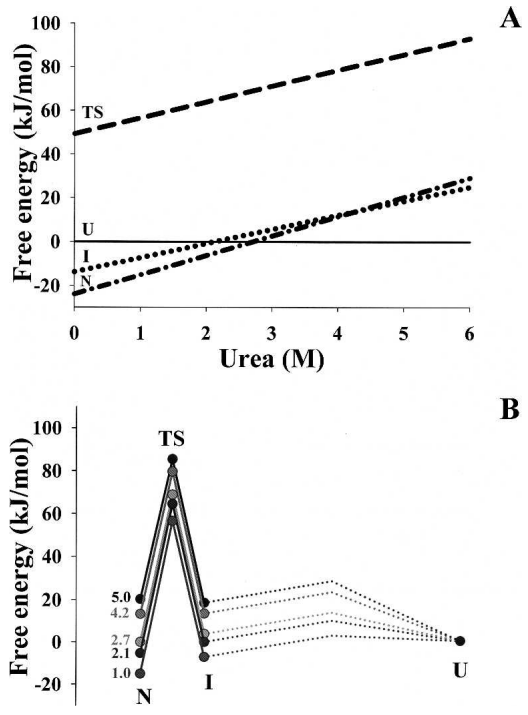

Figure 8. Graphical presentations of the energy levels. (*A*) Urea concentration dependence of free energy levels of the three (N, I, U) conformational states and the main transition state (TS) of apomyoglobin. Free energy levels are calculated from Equations 7, 9, and 12; all free energies are counted off that of the unfolded state. The transmission coefficient κ is assumed to be equal to 1.0 (Fersht 2000). (*B*) A schematic presentation of free energy changing during apomyoglobin refolding/unfolding at various final urea concentrations indicated by numbers near the curves.

Table 1. *Thermodynamic characteristics of transitions between the N, I, and U states estimated from equilibrium and kinetic data*

		Parameter Value	
Transition	Parameter	Kinetic	Equilibrium
I↔U	$\Delta G^0_{I\text{-}U}$ (kcal/mol)	3.7 ± 0.5	2.6 ± 1
	$m_{I\text{-}U}$ (kcal/mol·[M])	1.7 ± 0.2	1.2 ± 0.4
	C_m (M)	2.2 ± 0.2	2.2 ± 0.5
N↔I	$\Delta G^0_{N\text{-}I}$ (kcal/mol)	3.4 ± 0.5	2.9 ± 1
	$m_{N\text{-}I}$ (kcal/mol·[M])	0.7 ± 0.2	1.0 ± 0.4
	C_m (M)	4.2 ± 0.2	3.6 ± 0.5
N↔U	$\Delta G^0_{N\text{-}U}$ (kcal/mol)	6.2 ± 0.5	6.2 ± 1
	$m_{N\text{-}U}$ (kcal/mol·[M])	2.3 ± 0.2	2.2 ± 0.4
	C_m (M)	2.8 ± 0.2	2.9 ± 0.5

$\Delta G^0_{I\text{-}U}$, $\Delta G^0_{N\text{-}I}$, $\Delta G^0_{N\text{-}U}$: free energies of the transitions I↔U, N↔I, and N↔U, respectively; $m_{I\text{-}U}$, $m_{N\text{-}I}$, $m_{N\text{-}U}$: urea-dependence coefficients for the transitions I↔U, N↔I, and N↔U, respectively; C_m: urea concentration at which the mid-transition between two states is realized.

The main thermodynamic characteristics of transitions between the N, I, and U states obtained from our kinetic and equilibrium experiments are presented in Table 1. The equilibrium parameters were derived from the three-state model (Tanford 1968; Pace 1986). The kinetic and equilibrium results are seen to be close. The error of equilibrium parameters exceeds that of kinetic ones due to a low population of the equilibrium intermediate, and, hence, to a larger error in separation of the individual transitions. In contrast, kinetic experiments allow a more precise separation of the transitions due to a large difference in their rates.

Knowledge of free energies of the U, I, N, and TS states allows structural characterization of the transition state by "Φ_{TS}-analysis" and that of the kinetic intermediate state by "Φ_I-analysis" using the protein engineering by site-directed

mutations (Matouschek et al. 1990, 1998; Itzhaki et al. 1995). The results of the current work allow estimating these free energies and can be useful in Φ-analysis of apomyoglobin that folds through an intermediate, and in studies on the transition state structure.

Materials and methods

Expression and isolation

Engineered sperm whale apomyoglobin was isolated and purified according to Jennings et al. (1995) after pET17a plasmid expression in *Escherichia coli* BL21 (DE3) cells. The protein purity was checked by SDS electrophoresis. Protein concentration was calculated from absorbance at 280 nm using the extinction coefficient $A_{280}^{0.1\%} = 0.8$ (Harrison and Blout 1965).

Circular dichroism

Far UV CD spectra of the protein at various urea concentrations were registered with a J-600 spectropolarimeter (Jasco) at a protein concentration of 1 mg/mL using a 0.1-mm pathway quartz cell.

Tryptophan (Trp) fluorescence

Experiments on equilibrium urea-induced protein unfolding were carried out using a Shimadzu RF-5301PC spectrofluorimeter. Measurements were taken in 10 mM Na-acetate buffer at pH 6.2, 11°C. The protein concentration was 0.03 mg/mL.

Kinetic measurements were taken using a home-made stopped-flow rapid mixing attachment developed in collaboration with Dr. T. Nagamura (Unisoku Inc., Hirakata, Osaka, Japan). Two pneumatic drive syringes of the volume ratio 1:6 with a mixer and a 20-μL measuring cell were mounted inside the temperature-controlled block with a temperature control precision of 0.1°C. The time constant (integration time) was 0.002 sec. The stopped-flow attachment was combined with a 150 W Xe-lamp light source (LOMO), excitation and emission monochromators (MDR-4, LOMO), and a recording system including a personal computer and an amplifier with a sufficient capability for varying the time constant. The final protein concentration was 0.03 mg/mL. The initial urea concentration was 5.0 M in experiments on protein refolding and 0.0 M in experiments on unfolding. Because apomyoglobin folds quite rapidly at room temperature, a lower temperature should be used to slow down the reaction rate. On the other hand, the possibility of cold denaturation of this protein previously reported by Griko et al. (1988) made us choose 11°C for our experiments (Baryshnikova et al. 2005).

Acknowledgments

The authors thank P.E. Wright for kindly providing apomyoglobin plasmid. We thank N.B. Ilyina and I.A. Kashparov for excellent technical assistance. This work was supported in part by HHMI award 55000305, by award RB2–2022 from the US Civilian Research & Development Foundation for the Independent States of the Former Soviet Union (CRDF, to E.N.B. and V.E.B.), and by RAS Programmes on MCB and for "Scientific schools."

References

Baldwin, R.L. 1996. On-pathway versus off-pathway folding intermediates. *Fold. & Des.* **1:** R1–R8.

Barrick, D. and Baldwin, R.L. 1993a. The molten globule intermediate of apomyoglobin and the process of protein folding. *Protein Sci.* **2:** 869–876.

———. 1993b. Three-state analysis of sperm whale apomyoglobin folding. *Biochemistry* **32:** 3790–3796.

Baryshnikova, E.N., Sharapov, M.G., Kashparov, I.A., Ilyina, N.B., and Bychkova, V.E. 2005. Investigation of apomyoglobin stability depending on urea and temperature at two different pH values. *Mol. Biol.* **39:** 292–297.

Burns, L.L., Dalessio, P.M., and Ropson, I.J. 1998. Folding mechanism of three structurally similar β-sheet proteins. *Proteins* **33:** 107–118.

Cavagnero, S., Dyson, H.J., and Wright, P.E. 1999. Effect of H helix destabilizing mutations on the kinetic and equilibrium folding of apomyoglobin. *J. Mol. Biol.* **285:** 269–282.

Clarke, A.R. and Waltho, J.P. 1997. Protein folding pathways and intermediates. *Curr. Opin. Biotechnol.* **8:** 400–410.

Dolgikh, D.A., Gilmanshin, R.I., Brazhnikov, E.V., Bychkova, V.E., Semisotnov, G.V., Venyaminov, S.Y., and Ptitsyn, O.B. 1981. α-Lactalbumin: Compact state with fluctuating tertiary structure? *FEBS Lett.* **136:** 311–315.

Eliezer, D. and Wright, P.E. 1996. Is apomyoglobin a molten globule? Structural characterization by NMR. *J. Mol. Biol.* **263:** 531–538.

Eliezer, D., Yao, J., Dyson, H.J., and Wright, P.E. 1998. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* **5:** 148–155.

Fersht, A.R. 2000. *Structure and mechanism in protein science*, 3rd ed. W.H. Freeman & Co., New York.

Fersht, A.R., Matouschek, A. and Serrano, L. 1992. The folding of an enzyme. 1. Theory of protein engineering analysis and pathway of protein folding. *J. Mol. Biol.* **224:** 771–782.

Garcia, C., Nishimura, C., Cavagnero, S., Dyson, H.J., and Wright, P.E. 2000. Changes in the apomyoglobin folding pathway caused by mutation of the distal histidine residue. *Biochemistry* **39:** 11227–11237.

Griko, Y.V., Privalov, P.L., Venyaminov, S.Y., and Kutyshenko, V.P. 1988. Thermodynamic study of the apomyoglobin structure. *J. Mol. Biol.* **202:** 127–138.

Harrison, S.C. and Blout, E.R. 1965. Reversible conformational changes of myoglobin and apomyoglobin. *J. Biol. Chem.* **61:** 623–627.

Hughson, F.M., Wright, P.E., and Baldwin, R.L. 1990. Structural characterization of a partly folded apomyoglobin intermediate. *Science* **249:** 1544–1548.

Itzhaki, L.S., Otzen, D.E., and Fersht, A.R. 1995. The structure of the transition state for folding of Chymotrypsin Inhibitor 2 analyzed by protein engineering methods: Evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254:** 260–288.

Jamin, M. and Baldwin, R.L. 1998. Two forms of the pH 4 folding intermediate of apomyoglobin. *J. Mol. Biol.* **276:** 491–504.

Jamin, M., Yeh, S.-R., Rousseau, D.L., and Baldwin, R.L. 1999. Submillisecond unfolding kinetics of apomyoglobin and its pH 4 intermediate. *J. Mol. Biol.* **292:** 731–740.

Jamin, M., Antalik, M., Loh, S.N., Bolen, D.W., and Baldwin, R.L. 2000. The unfolding enthalpy of the pH 4 molten globule of apomyoglobin measured by isothermal titration calorimetry. *Protein Sci.* **9:** 1340–1346.

Jennings, P.A. and Wright, P.E. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **262:** 892–896.

Jennings, P.A., Stone, M.J., and Wright, P.E. 1995. Overexpression of myoglobin and assignment of its amide, Cα and Cβ resonances. *J. Biomol. NMR* **6:** 271–276.

Kay, M.S. and Baldwin, R.L. 1996. Packing interactions in apomyoglobin folding intermediate. *Nat. Struct. Biol.* **3:** 439–445.

Kim, P.S. and Baldwin, R.L. 1990. Intermediates in the folding reactions of small proteins. *Annu. Rev. Biochem.* **59:** 631–666.

Laurents, D.V., Corrales, S., Elias-Arnanz, M., Sevilla, P., Rico, M., and Padmanabhan, S. 2000. Folding kinetics of phage 434 Cro protein. *Biochemistry* **39:** 13963–13973.

Lecomte, J.T., Sukits, S.F., Bhattacharya, S., and Falzone, C.J. 1999. Conformational properties of native sperm whale apomyoglobin in solution. *Protein Sci.* **8:** 1484–1491.

Matouschek, A., Kellis Jr., J.T., Serrano, L., and Fersht, A.R. 1990. Transient folding intermediates characterized by protein engineering. *Nature* **346:** 440–445.

———. 1998. Mapping the transition state and pathway of protein folding by protein engineering. *Nature* **340:** 122–126.

Munoz, V., Lopez, E.M., Jager, M., and Serrano, L. 1994. Kinetic characterization of the chemotactic protein from *Escherichia coli*, CheY. Kinetic analysis of the inverse hydrophobic effect. *Biochemistry* **33:** 5858–5866.

Pace, C.N. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol*. **131:** 266–280.

Parker, M.J. and Marqusee, S. 1999. The cooperativity of burst phase reactions explored. *J. Mol. Biol*. **293:** 1195–1210.

Parker, M.J., Spencer, J., and Clarke, A.R. 1995. An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol*. **253:** 771–786.

Parker, M.J., Dempsey, C.E, Lorch, M., and Clarke, A.R. 1997. Acquisition of native β-strand topology during the rapid collapse phase of protein folding. *Biochemistry* **36:** 13396–13405.

Ptitsyn, O.B. 1995. Molten globule and protein folding. *Adv. Protein Chem*. **47:** 83–229.

Roder, H. and Colon, W. 1997. Kinetic role of early intermediates in protein folding. *Curr. Opin. Struct. Biol*. **7:** 15–28.

Santoro, M.M. and Bolen, D.W. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry* **27:** 8063–8068.

Schreiber, G. and Fersht, A.R. 1993. The refolding of *cis*- and *trans*-peptidylprolyl isomers of barstar. *Biochemistry* **32:** 11195–11203.

Tanford, C. 1968. Protein denaturation. Part B. The transition from native to denatured state. *Adv. Protein Chem*. **23:** 218–275.

Tang, K.S, Guralnick, B.J., Wang, W.K., Fersht, A.R., and Itzhaki, L.S. 1999. Stability and folding of the tumour suppressor protein p16. *J. Mol. Biol*. **285:** 1869–1886.

Tcherkasskaya, O. and Ptitsyn, O.B. 1999. Direct energy transfer to study the 3D structure of non-native proteins: AGH complex in molten globule state of apomyoglobin. *Protein Eng*. **12:** 485–490.

Tcherkasskaya, O., Bychkova, V.E., Uversky, V.N., and Gronenborn, A.M. 2000. Multisite fluorescence in proteins with multiple tryptophan residues. Apomyoglobin natural variants and site-directed mutants. *J. Biol. Chem*. **275:** 36285–36294.

Ternstorm, T., Mayor, U., Akke, M., and Oliveberg, M. 1999. From snapshot to movie: φ-analysis of protein folding transition states taken one step further. *Proc. Natl. Acad. Sci*. **96:** 14854–14859.

Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, H.J., and Wright, P.E. 1999. Quench-flow experiments combined with mass spectrometry show apomyoglobin folds through an obligatory intermediate. *Protein Sci*. **8:** 45–49.